(19) **United States**

(12) **Patent Application Publication** (10) Pub. No.: **US 2015/0293772 A1**

GANESAN et al. (43) **Pub. Date:** **Oct. 15, 2015**

(54) **VIRTUAL SWITCH AND VIRTUAL SWITCH PORT MANAGEMENT FOR VM AVAILABILITY**

(71) Applicant: **VMWARE, INC.**, Palo Alto, CA (US)

(72) Inventors: **Hariharan Jeyaraman GANESAN**, Bangalore (IN); **Jinto ANTONY**, Bangalore (IN)

(73) Assignee: **VMWARE, INC.**, Palo Alto, CA (US)

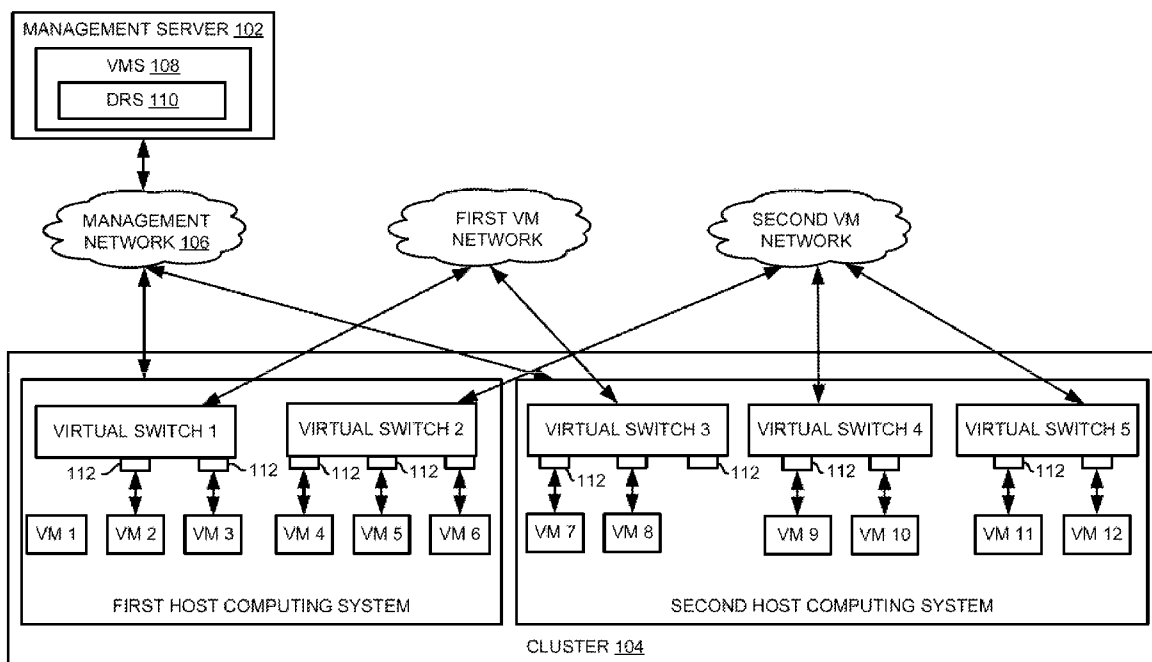(21) Appl. No.: **14/285,642**

(22) Filed: **May 23, 2014**

(30) **Foreign Application Priority Data**

Apr. 11, 2014 (IN) 1934/CHE/2014

Publication Classification

(57) **ABSTRACT**

Techniques for virtual switch and virtual switch port management for VM availability in a cluster are described. In one example embodiment, a determination is made as to whether a virtual switch port on a first virtual switch associated with a first VM network is available for powering on the VM on a first host computing system. Based on the outcome of the determination either further determination is then made as to whether a virtual switch port on a second virtual switch associated with the first VM network is available to power on the VM on a second host computing system or migration of the VM in a power-off state is initiated to the second host computing system and powered-on on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network associated.
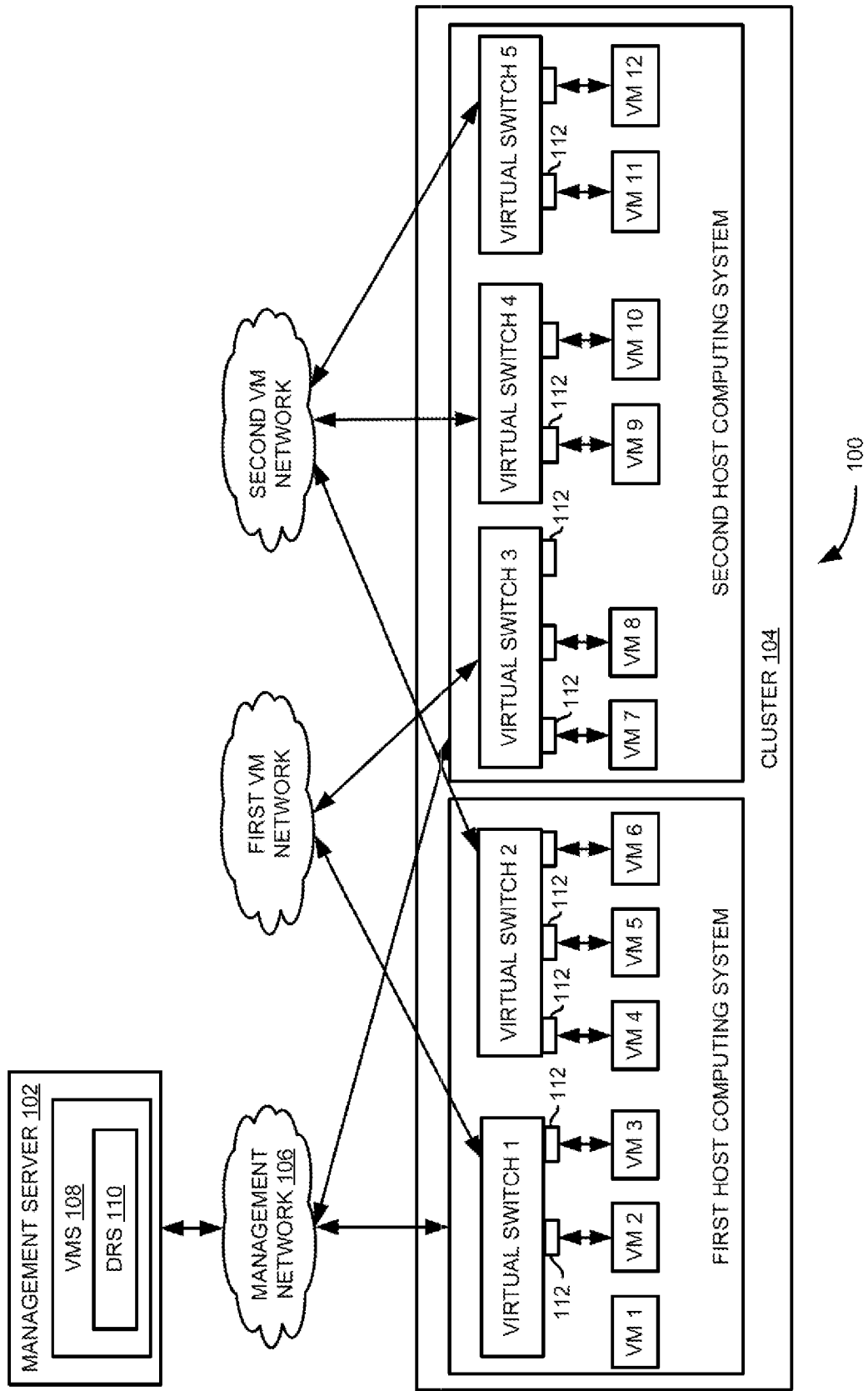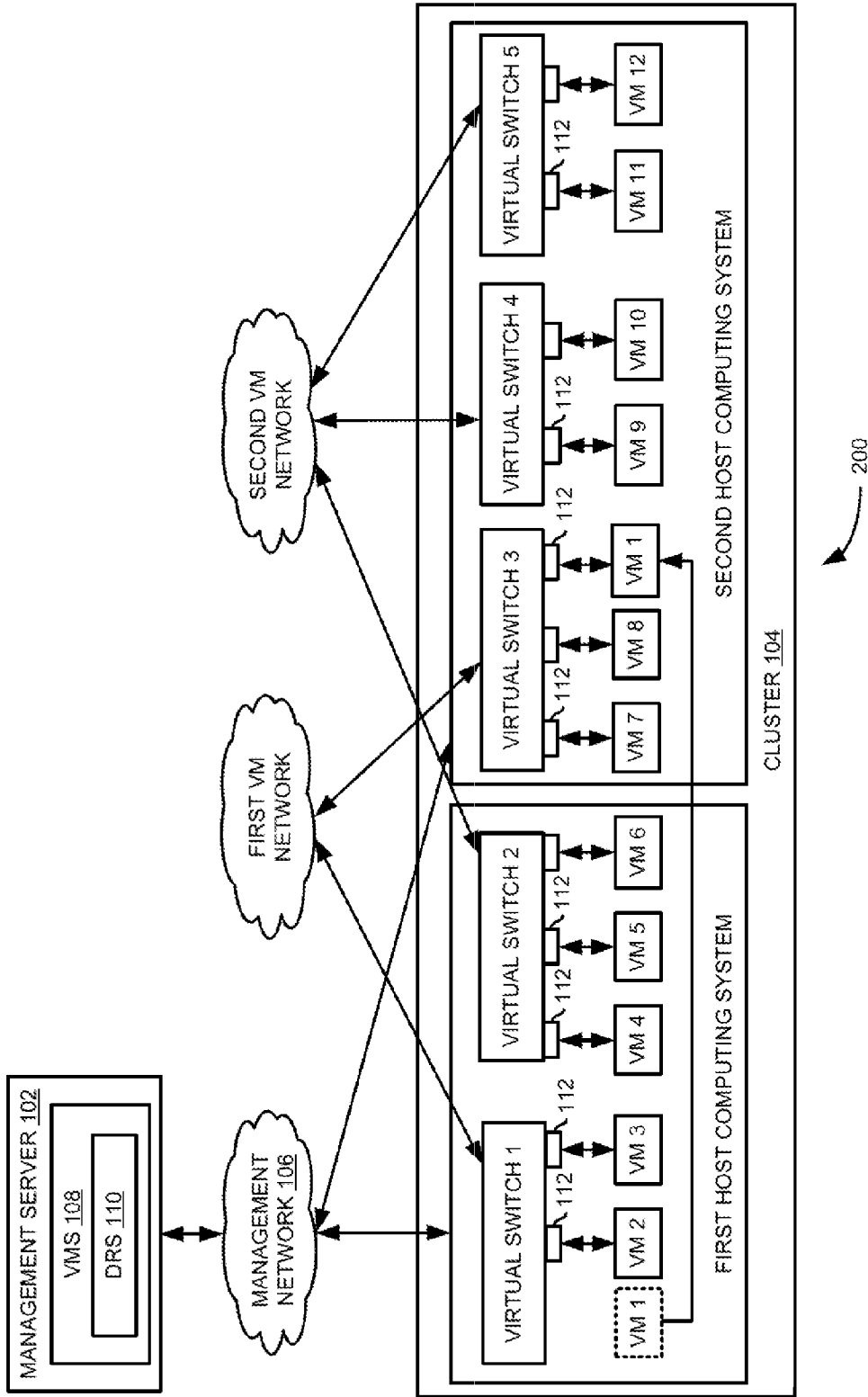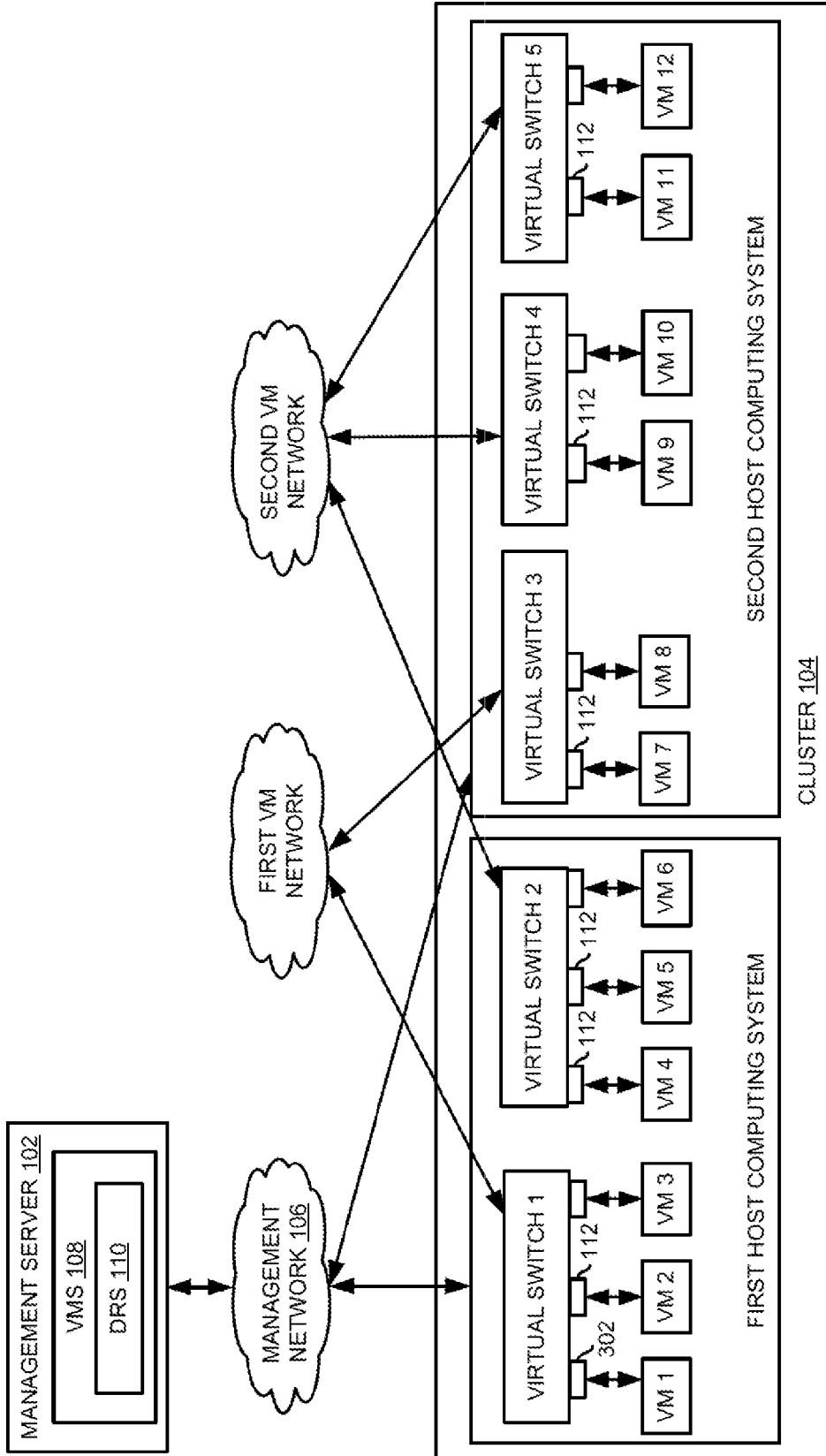
FIG. 1

FIG. 2

FIG. 3

402

IS A VIRTUAL SWITCH PORT ON A FIRST VIRTUAL SWITCH ASSOCIATED WITH A FIRST VM NETWORK AVAILABLE FOR POWERING ON A VM ON A FIRST HOST COMPUTING SYSTEM?

YES → 404

POWER ON THE VM ON THE FIRST HOST COMPUTING SYSTEM

NO

406

IS A VIRTUAL SWITCH PORT ON A SECOND VIRTUAL SWITCH ASSOCIATED WITH THE FIRST VM NETWORK AVAILABLE TO POWER ON THE VM ON A SECOND HOST COMPUTING SYSTEM?

YES → 410

INITIATE MIGRATION OF THE VM IN A POWER-OFF STATE TO THE SECOND HOST COMPUTING SYSTEM AND POWER ON THE VM ON THE SECOND HOST COMPUTING SYSTEM VIA THE VIRTUAL SWITCH PORT ON THE SECOND VIRTUAL SWITCH ASSOCIATED WITH THE FIRST VM NETWORK

NO → 408

INSTRUCT THE FIRST VIRTUAL SWITCH ASSOCIATED WITH THE FIRST HOST COMPUTING SYSTEM TO DYNAMICALLY CREATE THE VIRTUAL SWITCH PORT THAT IS ASSOCIATED WITH THE FIRST VM NETWORK FOR POWERING ON THE VM

400

FIG. 4

START ~ 502

504

ANY VM POWER ON
OPERATION?

NO

YES

506

IS A VIRTUAL SWITCH
PORT AVAILABLE ON A
VIRTUAL SWITCH TO
POWER ON A VM?

YES

NO

508

SEARCH ANY OTHER HOST COMPUTING SYSTEM
HAVING A UNUSED VIRTUAL SWITCH PORT ON A
VIRTUAL SWITCH CONNECTED TO SAME VM NETWORK

510

IS A VIRTUAL SWITCH
PORT AVAILABLE ON ANY
OTHER HOST
COMPUTING SYSTEM?

NO

YES

514

RECONFIGURE THE VIRTUAL
SWITCH AND TO ADD A
VIRTUAL SWITCH PORT TO
ACCOMMODATE THE NEWLY
POWERED ON VM

512

COLD MIGRATE THE
VM TO THE IDENTIFIED
HOST COMPUTING
SYSTEM

516

POWER ON THE VM

500
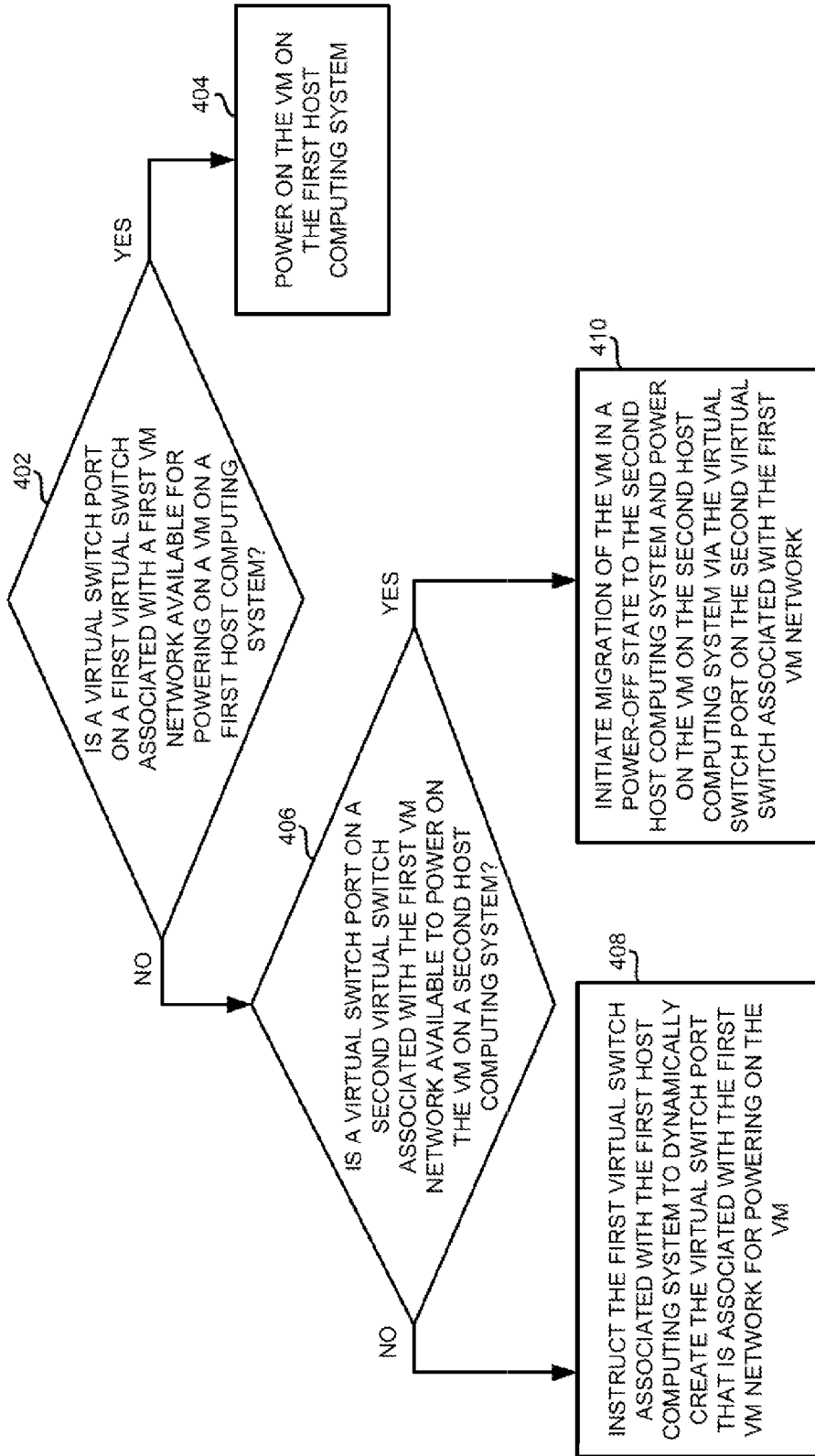
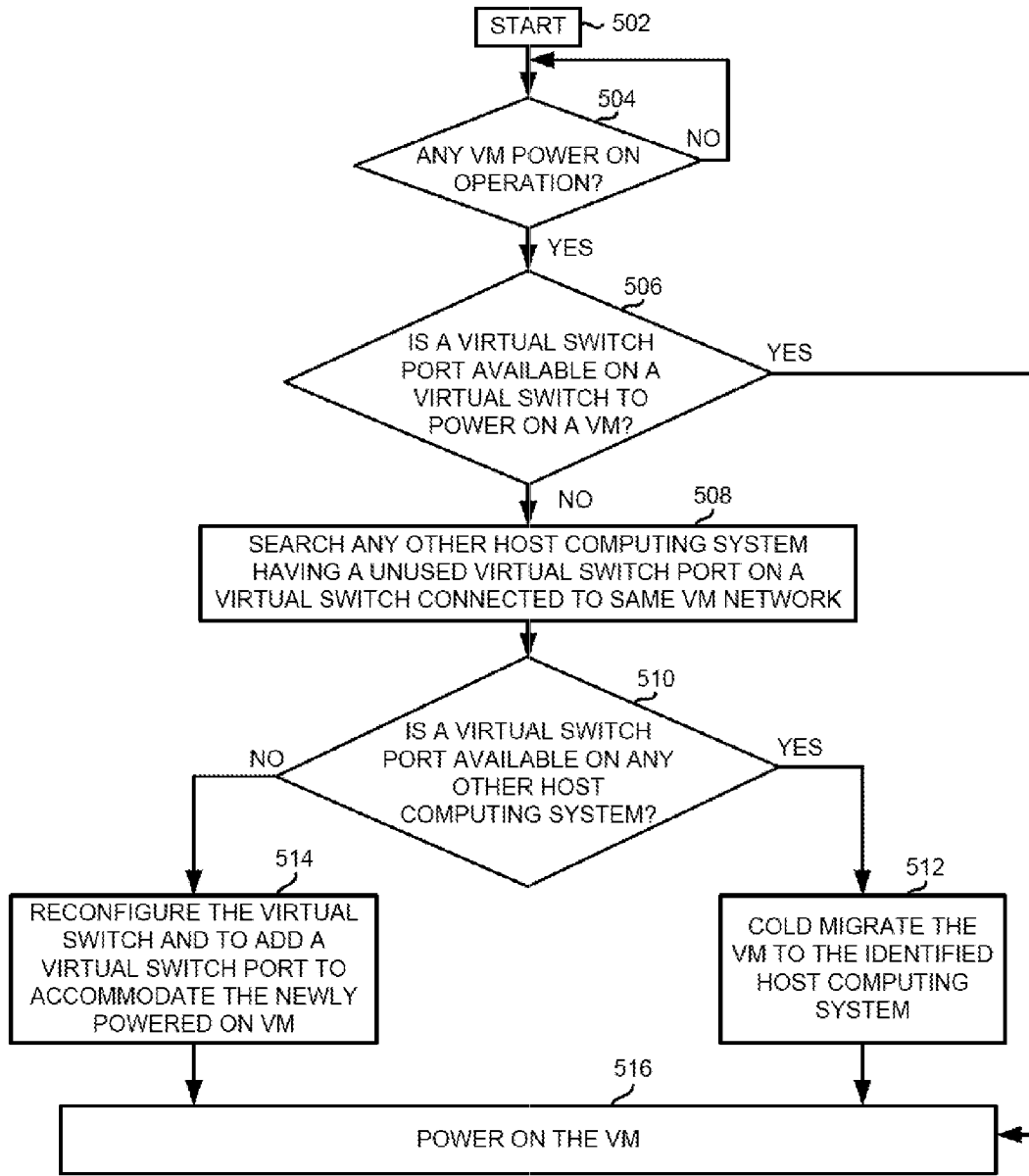FIG. 5

VIRTUAL SWITCH AND VIRTUAL SWITCH PORT MANAGEMENT FOR VM AVAILABILITY

RELATED APPLICATIONS

[0001] Benefit is claimed under 35 U.S.C. 119(a)-(d) to Foreign application Serial No. 1934/CHE/2014 filed in India entitled "VIRTUAL SWITCH AND VIRTUAL SWITCH PORT MANAGEMENT FOR VM AVAILABILITY", filed on Apr. 11, 2014, by VMware, Inc., which is herein incorporated in its entirety by reference for all purposes.

BACKGROUND

[0002] Virtual networking can network virtual machines (VMs) in a manner analogous to the way physical networking is used to network physical devices. Typical virtual networking components include virtual switches, virtual switch ports, and virtual network interface controllers (vNICs). Virtual switches allow VMs on the same physical host computing system, or across physical host computing systems, to communicate with each other using the same protocols that would be used over physical switches, without the need for additional networking hardware. Multiple virtual switches can be created on a single physical host computing system. Virtual switch ports on a virtual switch provide logical connection points for virtual devices, such as VMs, and between virtual and physical devices. Typically, each virtual switch is capable of having several thousand virtual switch ports.

[0003] In most cases, a virtual switch port must be available for a VM to power on. Typically, administers will configure a virtual switch to have enough virtual switch ports to accommodate the maximum anticipated number of VMs. However, this may result in consuming additional resources if a separate process is used for implementing each virtual switch. Further, this can result in a significant number of unused virtual switch ports, which can introduce security threats and expose VMs to external attacks and exploits, such as an intruder gaining unauthorized access, malware, spam and/or phishing.

SUMMARY

[0004] One or more embodiments disclosed herein provide a method for virtual switch and virtual switch port management for VM availability in a virtual datacenter. The method includes determining whether a virtual switch port on a first virtual switch associated with a first VM network is available for powering on the VM on a first host computing system. The method further determines whether a virtual switch port on a second virtual switch associated with the first VM network is available to power on the VM on a second host computing system, if the virtual switch port on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system. migration of the VM in a power-off state is then initiated to the second host computing system and powering on the VM on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network, if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system.

[0005] Further embodiments of the present disclosure include a non-transitory computer-readable storage medium that includes instructions that enable a processing unit to implement one or more of the methods set forth above or the functions of the computer system set forth above. In one embodiment, a non-transitory computer-readable storage medium is provided having instructions that manage execution of a virtual machine. The instructions, when executed in a computing device, perform the steps for virtual switch and virtual switch port management for VM availability in a virtual datacenter. The instructions further include determining whether a virtual switch port on a first virtual switch associated with a first VM network is available for powering on the VM on a first host computing system. The instruction furthermore includes determining whether a virtual switch port on a second virtual switch associated with the first VM network is available to power on the VM on a second host computing system if the virtual switch port on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system. Moreover the instruction initiate migration of the VM in a power-off state to the second host computing system and powering on the VM on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network, if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system.

[0006] Embodiments of the present disclosure provide a computing system. The computing system includes first and second host computing systems in a virtual datacenter. Each of the host computing include multiple virtual switches, Wherein each virtual switch having multiple virtual switch ports and wherein each host computing system hosting multiple VMs via the associated multiple virtual switches and multiple switch ports. The computing system further includes a management network that is communicatively coupled to the first and second host computing systems. In addition, the computing system includes first and second VM networks communicatively coupled to the associated multiple virtual switches residing in the first and second host computing systems for routing network traffic of VMs. Moreover, the computing system includes a management server that is communicatively to the management network, wherein the management server includes a dynamic resource scheduler (DRS), wherein the DRS is configured for virtual switch and virtual switch port management for VM availability in the virtual datacenter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIGS. 1-3 are block diagrams illustrating a system for virtual switch and virtual switch port management for VM in a virtual datacenter, according to an example embodiment.

[0008] FIG. 4 is a flow diagram of process for virtual switch and virtual switch port management for VM in a virtual datacenter, according to an example embodiment.

[0009] FIG. 5 is another flow diagram of process for virtual switch and virtual switch port management for VM in a virtual datacenter, according to an example embodiment.

DETAILED DESCRIPTION

[0010] Embodiments described herein provide virtual switch and virtual switch port management for VM availability in a virtual datacenter. The proposed technique enables powering on VMs via cold migration or dynamic virtual port switch configuration even when virtual switch ports are not

available to power on VMs in an associated virtual switch. Further, this technique significantly reduces memory utilization for virtual networking. Furthermore, this technique enables improved security by not having free virtual switch ports, i.e., having more virtual switch ports than needed in virtual network, which can be prone for external attacks.

System Overview and Examples of Operation

[0011] FIG. 1 is a block diagram illustrating system 100 for providing virtual switch and virtual switch port management for VM availability in a cluster in a virtual datacenter, according to an example embodiment. As shown in FIG. 1, system 100 includes first and second host computing systems and associated virtual machines (VMs) VM1-12 hosted by first and second host computing systems via associated virtual switches 1-5 and virtual switch ports 112 in a cluster 104. Further as shown in FIG. 1, system 100 includes management server 102 that is communicatively coupled to first and second host computing systems via management network 106. Furthermore as shown in FIG. 1, management server 102 includes dynamic resource scheduler (DRS) 110 residing in virtual management software (VMS) 108. In addition as shown in FIG. 1, first and second host computing systems are communicatively coupled to associated first and second VM networks via associated virtual switches 1-5.

[0012] In operation, VMS 108 determines whether a virtual switch port 112 on a first virtual switch associated with a first VM network is available for powering on a VM (VM1 as shown in FIG. 1) on first host computing system. In some embodiments, DRS 110 residing in VMS 108 determines whether a virtual switch port 112 on a first virtual switch associated with a first VM network is available for powering on a VM (VM1 as shown in FIG. 1) on first host computing system.

[0013] Based on the outcome of the determination, VMS 108 powers on the VM (VM1 shown in FIG. 1) on first host computing system if the virtual switch port 112 on the first virtual switch associated with the first VM network is available for powering on the VM (VM1 shown in FIG. 1) on the first host computing system. Further based on the outcome of the determination, VMS 108 furthermore determines whether a virtual switch port 112 on a second virtual switch associated with the first VM network is available to power on the VM (VM1 shown in FIG. 1) on a second host computing system if the virtual switch port 112 on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system.

[0014] In some embodiments, based on the outcome of the determination, DRS 110 powers on the VM (VM1 shown in FIG. 1) on first host computing system if the virtual switch port 112 on the first virtual switch associated with the first VM network is available for powering on the VM (VM1 shown in FIG. 1) on the first host computing system. Further based on the outcome of the determination, DRS 110 furthermore determines whether a virtual switch port 112 on a second virtual switch associated with the first VM network is available to power on the VM (VM1 shown in FIGS. 1-3) on a second host computing system if the virtual switch port 112 on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system.

[0015] Further in operation, as shown in FIG. 2, based on the outcome of the determination,. VMS 108 initiates migration of the VM (VM 1 shown in FIG. 2) in a power-off state

(also referred to as cold migration) to second host computing system and powers on the VM (VM1 shown in FIG. 2) on second host computing system via virtual switch port 112 on second virtual switch associated with first VM network if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system. One skilled in the art would appreciate that by giving VM migration a priority over creating a virtual switch port may reduce open virtual ports in the network which could help reduce security risks as any open virtual switch port may be used by any malicious software program or process to initiate external attacks, such as malware, span and/or phishing.

[0016] In some embodiments, as shown in FIG. 2, based on the outcome of the determination, DRS 110 initiates migration of the VM (VM1 shown in FIG. 2) in a power-off state (also referred to as cold migration) to second host computing system and powers on the VM (VM1 shown in FIG. 2) on second host computing system via virtual switch port 112 on second virtual switch associated with first VM network if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system.

[0017] Furthermore in operation, as shown in FIG. 3, VMS 108 instructs first virtual switch associated with first host computing system to dynamically create a virtual switch port 302 associated with first VM network for powering on the VM (VM1 shown in FIGS. 1-3) if the virtual switch port 112 on second virtual switch associated with first VM network is not available to power on the VM (VM1 shown in FIGS. 1-3) on second host computing system.

[0018] In some embodiments, as shown in FIG. 3, DRS 110 instructs first virtual switch associated with first host computing system to dynamically create a virtual switch port 302. associated with first VM network for powering on the VM (VM1 shown in FIGS. 1-3) if the virtual switch port 112 on second virtual switch associated with first VM network is not available to power on the VM (VM1 shown in FIGS. 1-3) on second host computing system.

[0019] Also, although certain terms are used primarily herein, other terms could be used interchangeably to yield equivalent embodiments and examples. For example, the term "host computing system" may be used interchangeably with "physical host", "physical machine" or "physical device". Further for example, there term "virtual switch port" may he used interchangeably with "virtual port". Furthermore for example, it is well-known that equivalent terms in the field of system virtualization or similar or related fields could be substituted for such terms as "physical computer," "hypervisor," "virtual machine," or the like. Also, the terms "virtual computing environment" and "virtual datacenter" are used interchangeably throughout the document. The terms "network failure", "network connectivity failure", and "lost network connectivity" are used interchangeably throughout the document.

[0020] Numerous specific details are set forth herein, such as data formats and code sequences and the like, in order to provide a thorough understanding of the described techniques. The embodiments described also can be practiced without some of the specific details described herein, or with other specific details, such as changes with respect to the ordering of the logic, different logic, different architectures, or the like. Thus, the scope of the techniques and/or functions described is not limited by the particular order, selection, or

decomposition of aspects described with reference to any particular routine, module, component, or the like.

Example Processes

[0021] FIG. 4 is a flow diagram of process **400**, for providing virtual switch and virtual switch port management for VM availability in a virtual datacenter, according to an example embodiment.

[0022] At block **402**, process **400** determines whether a virtual switch port on a first virtual switch associated with a first VM network is available for powering on the VM on a first host computing system.

[0023] At block **404**, the process **400** powers on the VM on the first host computing system, if the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system. At **406**, the process **400** determines whether a virtual switch port on a second virtual switch associated with the first VM network is available to power on the VM on a second host computing system, if the virtual switch port on the first virtual switch associated with the first VII network is not available for powering on the VM on the first host computing system.

[0024] At block **408**, the first virtual switch associated with the first host computing system is instructed to dynamically create the virtual switch port that is associated with the first VM network for powering on the VM if the virtual switch port on the second virtual switch associated with the first VM network is not available to power on the VM on the second host computing system. At block **410**, migration of the VM in a power-off state is initiated to the second host computing system and powering on the VM on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network, if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system.

[0025] Process **400** for providing virtual switch and virtual switch port management for VM availability in a virtual datacenter is explained in more detail above with reference to the system diagrams **100-300** shown in FIGS. **1-3**.

[0026] FIG. **5** is a flow diagram of another process **500**, for providing virtual switch and virtual switch port management for VM availability in a virtual datacenter, according to an example embodiment.

[0027] At blocks **502** and **504**, process **500** starts with determining whether any VM is being powered on in the virtual datacenter. At block **506**, a determination is made as to whether a virtual switch port is available on a virtual switch to power on the VM. If a virtual switch port is available then the process **500** goes to block **516** and powers on the VM using the virtual switch port via associated virtual switch. If the virtual switch port is not available to power on the VM, then process at blocks **508** and **510** initiates a search to determine whether any other host computing system in the virtual datacenter has an unused virtual switch port available that is running on the associated virtual switch connected to the same VM network. At block **514**, the virtual switch is reconfigured to add a virtual switch port to accommodate powering on the VM and then goes to block **516** to power on the VM. At block **512**, a cold migration is performed to migrate the VM to the identified computing system and then goes to block **516** to power on the VM.

[0028] The architecture shown in FIGS. **1-3** may in some embodiments be partially or fully virtualized. For example,

systems **100-300** shown in FIGS. **1-3** may be one or possibly many VMs executing on physical hardware and managed by a hypervisor, VM monitor, or similar technology. Also, multiple host computing systems **106** A-N show in FIGS. **1-3** may include virtualization logic to manage multiple VMs.

[0029] In an example embodiment, components/modules of DRS are implemented using standard programming techniques. In other embodiments, DRS may be implemented as instructions processed by a VM that executes as one of other programs.

[0030] Furthermore, in some embodiments, some or all of the components of DRS may be implemented or provided in other manners, such as at least partially in firmware and/or hardware, including, but not limited to one or more application-specific integrated circuits ("ASICs"), standard integrated circuits, controllers executing appropriate instructions, and including microcontrollers and/or embedded controllers, field-programmable gate arrays ("FPGAs"), complex programmable logic devices ("CPLDs"), and the like. Some or all of the system components and/or data structures may also be stored as contents (e.g., as executable or other machine-readable software instructions or structured data) on a computer-readable medium (e.g., as a hard disk; a memory; a computer network or cellular wireless network or other data. transmission medium; or a portable media article to be read by an appropriate drive or via an appropriate connection, such as a DVD or flash memory device) so as to enable or configure the computer-readable medium and/or one or more associated computing systems or devices to execute or otherwise use or provide the contents to perform at least sonic of the described techniques.

[0031] Further, from the foregoing it will be appreciated that, although specific embodiments have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of this disclosure. For example, the methods, techniques, and systems for migrating \\$^7$Ms across host computing systems based on cache characteristics in a virtualized datacenter are applicable to other architectures or in other settings. For example, the described techniques may be employed as part of a cloud-based computing resource offering, wherein customers may pay to have higher importance levels associated with their activities, in order to obtain higher levels of service or availability. As another example, the described techniques may be employed to allocate resources or schedule CPU time at the process level within an operating system. Also, the methods, techniques, and systems discussed herein are applicable to differing protocols, communication media (optical, wireless, cable, etc.) and devices e.g., desktop computers, wireless handsets, electronic organizers, personal digital assistants, tablet computers, portable email machines, game machines, pagers, navigation devices, etc.).

1. A method. for virtual switch and virtual switch port management for virtual machine (VW availability in a cluster in a virtual datacenter, comprising:

determining whether a virtual switch port on a first virtual switch associated with a first. VM network is available for powering on the VM on a first host computing system;

if the virtual switch port on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system, determining whether a virtual switch port on a second

virtual switch associated with the first VM network is available to power on the VM on a second host computing system; and

if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system, initiating migration of the VM in a power-off state to the second host computing system and powering on the VM on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network.

2. The method of claim 1, further comprising:

if the virtual switch port on the second virtual switch associated with the first VM network is not available to power on the VM on the second host computing system, instructing the first virtual switch associated with the first host computing system to dynamically create the virtual switch port that is associated with the first VM network for powering on the VM.

3. The method of claim 1, further comprising:

if the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system, then powering on the VM on the first host computing system.

4. The method of claim 1, wherein determining whether the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system, comprises:

determining whether the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system by a dynamic resource scheduler (DRS).

5. The method of claim 4, wherein determining whether the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system, comprises:

determining whether the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system by the DRS.

6. A non-transitory computer-readable storage medium including instructions that, when executed in a computing system, to perform virtual switch and virtual switch port management for VM availability in a cluster in a virtual datacenter, the method comprising:

determining whether a virtual switch port on a first virtual switch associated with a first VM network is available for powering on the VM on a first host computing system;

if the virtual switch port on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system, determining whether a virtual switch port on a second virtual switch associated with the first VM network is available to power on the VM on a second host computing system; and

if the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system, initiating migration of the VM in a power-off state to the second host computing system and powering on the VM on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network.

7. The non-transitory computer-readable storage medium of claim 6, further comprising:

if the virtual switch port on the second virtual switch associated with the first VM network is not available to power on the VM on the second host computing system, instructing the first virtual switch associated with the first host computing system to dynamically create the virtual switch port that is associated with the first VM network for powering on the VM.

8. The non-transitory computer-readable storage medium of claim 6, further comprising:

if the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system, then powering on the VM on the first host computing system.

9. The non-transitory computer-readable storage medium of claim 6, wherein determining whether the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system, comprises:

determining whether the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system by a dynamic resource scheduler (DRS).

10. The non-transitory computer-readable storage medium of claim 9, wherein determining whether the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system, comprises:

determining whether the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system by the DRS.

11. A computing system for virtual switch and virtual switch port management for VM availability in a cluster in a virtual datacenter, the system comprising:

first and second host computing systems, wherein each host computing comprises:

multiple virtual switches and wherein each virtual switch having multiple virtual switch ports and wherein each host computing system hosting multiple VMs via the associated multiple virtual switches and multiple switch ports;

a management network communicatively coupled to the first and second host computing systems;

first and second VM networks communicatively coupled to the associated multiple virtual switches residing in the first and second host computing systems for routing network traffic of VMs; and

a management server communicatively coupled to the management network, wherein the management server comprising virtual management software (VMS), and wherein the VMS comprising a dynamic resource scheduler (DRS), and they are configured to:

determine whether a virtual switch port on a first virtual switch associated with a first VM network is available for powering on the VM on a first host computing system;

if the virtual switch port on the first virtual switch associated with the first VM network is not available for powering on the VM on the first host computing system, determine whether a virtual switch port on a second virtual switch associated with the first VM

network is available to power on the VM on a second host computing system; and

if the virtual switch port on the second virtual switch associated With the first VM network is available to power on the VM on the second host computing system, initiate migration of the VM in a power-off state to the second host computing system and powering on the VM on the second host computing system via the virtual switch port on the second virtual switch associated with the first VM network.

12. The computing system of claim 11. further configured to:

if the virtual switch port on the second virtual switch associated with the first VM network is not available to power on the VM on the second host computing system, instruct the first virtual switch associated with the first host computing system to dynamically create the virtual switch port that is associated with the first VM network for powering on the VM.

13. The computing system of claim 11, further configured to:

if the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system, then powering on the VM on the first host computing system.

14. The computing system of claim 11, wherein determining whether the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system, comprises:

determining whether the virtual switch port on the first virtual switch associated with the first VM network is available for powering on the VM on the first host computing system by a dynamic resource scheduler (DRS).

15. The computing system of claim 14, Wherein determining whether the virtual switch port on the second virtual switch associated pith the first VM network is available to power on the VM on the second host computing system, comprises:

determining whether the virtual switch port on the second virtual switch associated with the first VM network is available to power on the VM on the second host computing system by the DRS.

* * * * *